

XP008008413

sub-

Q-32. Detection of Dechlorinating Bacteria in Groundwater and Soils from Waste Sites Contaminated with PCE and TCE

M.A. ELBERSON¹, R.C. EBERSOLE¹, R.M. YOUNG², J.A. TABINOWSKI¹, M.P. PERRY², S. FAHNESTOCK², E.R. HENDRICKSON¹,
¹Microbial and Industrial Microbiology, Life Science Enterprise, Du Pont Co., Glasgow Site, Newark, DE; ²Du Pont Co., Experimental Station, Wilmington, DE

Chlorinated ethenes, perchloroethene (PCE) and trichloroethene (TCE), are contaminants of groundwater at many hazardous waste sites. These solvents are both toxic and suspected carcinogens. Recently, *Dehalococcoides ethenogenes*, an organism described Maymo-Gatell et al. (Science 276, 1568-1571 (1997)), was shown to dechlorinate PCE and TCE by removing all the chlorine atoms, through a process known as dehalorespiration, to form ethane. In addition, other microorganisms, such as, *Dehalobacter restrictus* and *Dehalospirillum multivorans*, have demonstrated partial dechlorination of chlorinated ethenes. Based on their 16S rRNA gene sequences, we have developed a specific PCR assays to detect each of these organisms. Using these assays, the presence of these dehalorespiring organisms has been demonstrated in microcosms inoculated with groundwater or soil from four waste sites contaminated with PCE and TCE. The PCR amplicons were sequenced to ver-



Creation date: 10-16-2003
Indexing Officer: MBEYENE - MELKAM BEYENE
Team: OIPEScanning
Dossier: 09548998

Legal Date: 02-09-2002

No.	Doccode	Number of pages
1	SRNT	34

Total number of pages: 34

Remarks:

Order of re-scan issued on



A DOCPHOENIX

APPL PARTS

IMIS _____
Internal Misc. Paper

LET. _____
Misc. Incoming Letter

371P _____
PCT Papers in a 371 Application

A... _____
Amendment Including Elections

ABST _____
Abstract

ADS _____
Application Data Sheet

AF/D _____
Affidavit or Exhibit Received

APPENDIX _____
Appendix

ARTIFACT _____
Artifact

BIB _____
Bib Data Sheet

CLM _____
Claim

COMPUTER _____
Computer Program Listing

CRFL _____
All CRF Papers for Backfile

DIST _____
Terminal Disclaimer Filed

DRW _____
Drawings

FOR _____
Foreign Reference

FRPR _____
Foreign Priority Papers

IDS _____
IDS Including 1449

NPL _____
Non-Patent Literature

OATH _____
Oath or Declaration

PET. _____
Petition

RETMAIL _____
Mail Returned by USPS

SEQLIST _____
Sequence Listing

SPEC _____
Specification

SPEC NO _____
Specification Not in English

TRNA _____
Transmittal New Application

OUTGOING

CTMS _____
Misc. Office Action

1449 _____
Signed 1449

892 _____
892

ABN _____
Abandonment

APDEC _____
Board of Appeals Decision

APEA _____
Examiner Answer

CTAV _____
Count Advisory Action

CTEQ _____
Count Ex parte Quayle

CTFR _____
Count Final Rejection

Duplex
02/09/02

CTNF _____
Count Non-Final

CTRS _____
Count Restriction

EXIN _____
Examiner Interview

M903 _____
DO/EO Acceptance

M905 _____
DO/EO Missing Requirement

NFDR _____
Formal Drawing Required

NOA _____
Notice of Allowance

PETDEC _____
Petition Decision

INCOMING

AP.B _____
Appeal Brief

C.AD _____
Change of Address

N/AP _____
Notice of Appeal

PA.. _____
Change in Power of Attorney

REM _____
Applicant Remarks in Amendment

XT/ _____
Extension of Time filed separate

Internal

SRNT _____
Examiner Search Notes

CLMPTO _____
PTO Prepared Complete Claim Set

ECBOX _____
Evidence Copy Box Identification

WCLM _____
Claim Worksheet

WFEE _____
Fee Worksheet

File Wrapper

FWCLM _____
File Wrapper Claim

IIFW _____
File Wrapper Issue Information

SRFW _____
File Wrapper Search Info

Hart, Edward

59984

From: Loeb, Bronwen
Sent: Friday, February 08, 2002 6:32 PM
To: Hart, Edward
Subject: sequence search request 09/548,998

Please search SEQ ID No. 8 from 09/548,998 in all the nucleotide databases except the pending databases.

Thank you.

Bronwen Loeb, PhD

AU 1636
703-605-1197
CM1 12D12
Mailbox 11E-12

Edward Hart
Technical Support
Seq. ID No. 8
09/548,998

2/9/02

AB5504

1-NA

<u>Seq ID No</u>	<u># nt</u>
1	24
2	1377
3	1378
4	1377
5	1377
6	1377
8	47
30	51
34	49

GenCore version 4.5
Copyright: (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 1344.17 Seconds

576.837 Million cell updates/sec

Title:	US-09-548-998A-8
Perfect score:	47
Score:	14

sequence: 1 tgtgtrtgcygacatawgt.....ytaaagccgyaagygttg 47

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

1:	gb_ba:*
2:	gb_htg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_sy:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_v1:*
15:	em_ba:*
16:	em_fun:*
17:	em_fun:*
18:	em_in:*
19:	em_om:*
20:	em_om:*
21:	em_ov:*
22:	em_pat:*
23:	em_ph:*
24:	em_pl:*
25:	em_ro:*
26:	em_sts:*
27:	em_sy:*
28:	em_un:*
29:	em_v1:*
30:	em_htg_inu:*
31:	em_htg_inu:*
32:	em_htg_rod:*
33:	em_htg_num:*
34:	em_htg_inv:*
35:	em_htg_rod:*
36:	em_htg_other:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.2	94.0	47	6	AX039541	AX039541 Sequence
2	44.2	94.0	1212	6	AX039535	AX039535 Sequence
3	44.2	94.0	1212	6	AX039537	AX039537 Sequence
4	44.2	94.0	1212	6	AX039538	AX039538 Sequence
5	44.2	94.0	1212	6	AX039539	AX039539 Sequence
6	44.2	94.0	1322	1	AF357918	AF357918 Sequence
7	44.2	94.0	1332	1	UBA249252	UBA249252 Sequence
8	44.2	94.0	1335	6	AX039536	AX039536 Sequence
9	44.2	94.0	1421	1	AF230641	AF230641 Sequence
10	44.2	94.0	1434	1	AF004928	AF004928 Sequence
11	44.2	94.0	1443	6	AX039540	AX039540 Sequence
12	26.4	56.2	87286	2	AC014436	AC014436 Sequence
13	26.4	56.2	170914	3	AC010883	AC010883 Sequence
14	26.4	56.2	185810	3	AC010995	AC010995 Sequence
15	26.4	56.2	298839	3	AE003492	AE003492 Sequence
16	24.6	52.3	318503	2	AC074166	AC074166 Sequence
17	24.4	51.9	3215	5	ORNGGLUD	L34080 Orectochromis
18	24.4	51.9	3215	5	ORNGGLUD	L34080 Orectochromis
19	24.4	51.9	3391	5	ORNGGLUD	L46356 Orectochromis
20	24.4	51.9	59843	8	AF000816	L34036 Orectochromis
21	24.2	51.5	2421	5	AF389400	AF000816 Drya sat
22	24.2	51.5	57473	8	AF004128	AF389400 Danilo rer
23	23.8	50.6	1355	1	AF005746	U41288 Aflomyces m
24	23.6	50.2	410	1	AF010041	AF005746 Uncultured
25	23.4	49.8	110000	2	LMFCH32_14	AF010041 Unidentif
26	23	48.9	184841	2	AC079226	Continuation (15 o
27	22.8	48.5	19585	3	AF155610	AC079226 Homo sapi
28	22.8	48.5	119184	1	D90909	AF155610 Polistes
29	22.8	48.5	170139	2	AC073737	D90909 Synecocyst
30	22.8	48.5	195470	2	AC079367	AC073737 Mus muscu
31	22.8	48.5	235411	2	AC084066	AC079367 Mus muscu
32	22.6	48.1	173271	2	AC023919	AC084066 Mus muscu
33	22.4	47.7	160628	2	AC091243	AC023919 Homo sapi
34	22.4	47.7	184457	1	AF079317	AC091243 Trypanoso
35	22.2	47.2	2557	8	SCYLD063C	AF079317 Sphingomo
36	22.2	47.2	142047	9	AC026708	SCYLD063C S. cerevisia
37	22.2	47.2	160025	5	AC018748	AC026708 Homo sapi
38	22.2	47.2	162712	2	AC015677	AC018748 Homo sapi
39	22.2	47.2	165428	2	AC001099	AC015677 Homo sapi
40	22.2	47.2	172120	2	AC001886	AC001099 Homo sapi
41	22.2	47.2	179391	2	AP001283	AC001886 Homo sapi
42	22.2	47.2	203450	2	AP002009	AP001283 Homo sapi
43	22.2	47.2	208921	2	AP001911	AP002009 Homo sapi
44	22	46.8	1420	1	AF027031	AP001911 Homo sapi
45	22	46.8	1420	1	AF027032	AF027031 Unidentif

unsure /note="CONSENSUS"
5 /note="R-A/G"
unsure 11 /note="Y-C/T"
unsure 18 /note="W-A/T"
unsure 21 /note="Y-C/T"
unsure 28 /note="Y-T/C"
unsure 37 /note="Y-T/C"
unsure 42 /note="Y-C/T"
BASE COUNT 9 a 6 c 15 g 10 t 7 others
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tttgttgccgacatawgttgctcaataagccgaaagggcttg 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 TGTGTGGGCTGACATATGTTGTTCACTAAAGCCGTAAGGCTTG 47

RESULT 2
AX039535 1212 bp DNA PAT 18-NOV-2000
LOCUS Sequence 2 from Patent W00063443.
ACCESSION AX039535
VERSION AX039535.1 GI:11229557
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;

REFERENCE 1 (bases 1 to 1212)
AUTHORS Hendrickson,E.R. and Ebersole,R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria
JOURNAL Patent: WO 0063443-A 2 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
source
1. .1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"
BASE COUNT 321 a 252 c 369 g 269 t 1 others
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tttgttgccgacatawgttgctcaataagccgaaagggcttg 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 140 TGTGTGGGCTGACATATGTTGTTCACTAAAGCCGTAAGGCTTG 186

RESULT 3
AX039537 1212 bp DNA PAT 18-NOV-2000
LOCUS Sequence 4 from Patent W00063443.
ACCESSION AX039537
VERSION AX039537.1 GI:11229559
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.

REFERENCE 1 (bases 1 to 1212)
AUTHORS Hendrickson,E.R. and Ebersole,R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria
JOURNAL Patent: WO 0063443-A 2 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. .1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"
BASE COUNT 318 a 253 c 372 g 269 t
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tttgttgccgacatawgttgctcaataagccgaaagggcttg 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 140 TGTGTGGGCTGACATATGTTGTTCACTAAAGCCGTAAGGCTTG 186

RESULT 4
AX039538 1212 bp DNA PAT 18-NOV-2000
LOCUS Sequence 5 from Patent W00063443.
ACCESSION AX039538
VERSION AX039538.1 GI:11229560
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;

REFERENCE 1 (bases 1 to 1212)
AUTHORS Hendrickson,E.R. and Ebersole,R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria
JOURNAL Patent: WO 0063443-A 5 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
source
1. .1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"
BASE COUNT 320 a 253 c 371 g 268 t
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tttgttgccgacatawgttgctcaataagccgaaagggcttg 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 140 TGTGTGGGCTGACATATGTTGTTCACTAAAGCCGTAAGGCTTG 186

RESULT 5
AX039539 1212 bp DNA PAT 18-NOV-2000
LOCUS Sequence 6 from Patent W00063443.
ACCESSION AX039539
VERSION AX039539.1 GI:11229561
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Hendrickson,E.R. and Ebersole,R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria
JOURNAL Patent: WO 0063443-A 6 26-OCT-2000;

FEATURES	E. I. DU PONT DE NEMOURS AND COMPANY (US)
source	Location/Qualifiers
	1. 1112
	/organism="Dehalococcoides ethenogenes"
	/db_xref="taxon:61435"
BASE COUNT	320 a 255 c 371 g 266 t
ORIGIN	

```

VERSION      AF230641.1  GI:11545272
KEYWORDS     bacterium CBDB1.
SOURCE       bacterium CBDB1.
ORGANISM     Bacteria; Green non-sulfur bacteria; Dehalococcoides group.

REFERENCE    1 (bases 1 to 1421)
AUTHORS     Adrian, L., Szewzyk, U., Wecke, J. and Goerisch, H.
TITLE       Bacterial denaturation with chlorinated benzenes
JOURNAL     Nature 408 (6812), 580-583 (2000)
MEDLINE     21003821
REFERENCE    2 (bases 1 to 1421)
AUTHORS     Adrian, L., Tran, T., Szewzyk, U. and Goerisch, H.
TITLE       Direct Submission
JOURNAL     Submitted (31-JAN-2000) FG Technische Biochemie, TU Berlin, Seestr.
            13, Berlin 13353, Germany
FEATURES     Location/Qualifiers
             source          1..1421
                        /organism="bacterium CBDB1"
                        /strain="CBDB1"
                        /db_xref="taxon:114835"
                        /note="tentatively related to Dehalococcoides ethenogenes"
                        <1..>1421
                        /product="16S ribosomal RNA"

rRNA
BASE COUNT   368 a      307 c      435 g      311 t
ORIGIN

Query Match      94.0%; Score 44.2; DB 1; Length 1421;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tttgttggcgcgcacatawgttggttcaytaagccggaagggcttg 47
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 TGTGATGGGCTGACATTAAGCGGTTCATTAAAGCCGCAAGGCTTG 186

RESULT 10
LOCUS       AF004928      1434 bp      DNA      BCT      07-APR-2000
DEFINITION  Dehalococcoides ethenogenes 16S ribosomal RNA gene, partial
            sequence.
ACCESSION   AF004928
VERSION     AF004928.2  GI:7524025
KEYWORDS    Dehalococcoides ethenogenes.
            Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
            Dehalococcoides.
            1 (bases 1 to 1434)
            Maymo-Gatelli, X., Chien, Y., Gossett, J.M. and Zinder, S.H.
            Isolation of a bacterium that reductively dechlorinates
            tetrachloroethene to ethene
JOURNAL     Science 276 (5318), 1568-1571 (1997)
MEDLINE     9731370
REFERENCE    2 (bases 1 to 1434)
AUTHORS     Chien, Y. and Zinder, S.H.
TITLE       Direct Submission
JOURNAL     Submitted (20-MAY-1997) Microbiology, Cornell University, Wing
            Hall, Ithaca, NY 14853, USA
REFERENCE    3 (bases 1 to 1434)
AUTHORS     Chien, Y. and Zinder, S.H.
TITLE       Direct Submission
JOURNAL     Submitted (07-APR-2000) Microbiology, Cornell University, Wing
            Hall, Ithaca, NY 14853, USA
REMARK      Sequence update by submitter
            On Apr 7, 2000 this sequence version replaced gi:2232215.
FEATURES     Location/Qualifiers
             source          1..1434
                        /organism="Dehalococcoides ethenogenes"
                        /strain="strain 195"
                        /db_xref="taxon:61435"
                        /note="tentative genus and species name; Eubacterium not
                        closely related to presently known phyla"

```

```

rRNA
BASE COUNT   376 a      304 c      440 g      313 t      1 others
ORIGIN

Query Match      94.0%; Score 44.2; DB 1; Length 1434;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tttgttggcgcgcacatawgttggttcaytaagccggaagggcttg 47
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 TGTGATGGGCTGACATTAAGCGGTTCATTAAAGCCGCAAGGCTTG 186

RESULT 11
LOCUS       AX039540      1443 bp      DNA      PAT      18-NOV-2000
DEFINITION  Sequence 7 from Patent W00063443.
ACCESSION   AX039540
VERSION     AX039540.1  GI:11229562
KEYWORDS    Dehalococcoides ethenogenes.
            Dehalococcoides ethenogenes.
            Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
            Dehalococcoides.
            1 (bases 1 to 1443)
            Hendrickson, E.R. and Ebersole, R.C.
            Nucleic acid fragments for the identification of dechlorinating ba
            cteria
            Patent: WO 0063443-A 7 26-OCT-2000;
            E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES     Location/Qualifiers
             source          1..1443
                        /organism="Dehalococcoides ethenogenes"
                        /db_xref="taxon:61435"

BASE COUNT   379 a      306 c      443 g      314 t      1 others
ORIGIN

Query Match      94.0%; Score 44.2; DB 6; Length 1443;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tttgttggcgcgcacatawgttggttcaytaagccggaagggcttg 47
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 TGTGATGGGCTGACATTAAGCGGTTCATTAAAGCCGCAAGGCTTG 186

RESULT 12
LOCUS       AC014436/c      87286 bp      DNA      HTG      16-NOV-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
            pieces.
ACCESSION   AC014436
VERSION     AC014436.1  GI:6436899
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Fruit fly.
            Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 87286)
            Adams, M. and Venter, J.C.
            Direct Submission
            Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
            This sequence was identified as CDM:10212830 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

```

FEATURES

source

Location/Qualifiers

1. 87286 /organism="Drosophila melanogaster"

BASE COUNT

25135 a 18013 c 18299 g 25839 t

ORIGIN

Query Match

Best Local Similarity 56.2%; Score 26.4; DB 2; Length 87286;

Matches 27; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY

9 gcygacatawtygttcaytaagccgaaagcgtt 46

db 53263 GCTTACATATGTTCTCATTAAGCCTTAAGAACCT 53226

RESULT 13

LOCUS

AC010843 170914 bp DNA

INX 12-MAY-2001

DEFINITION Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone

AC010843 complete sequence.

ACCESSION

AC010843

VERSION

AC010843.10 GI:14028974

KEYWORDS

HTG.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 170914)

Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,

Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,

Rogers,Y., An,H., Baldwin,D., Banazon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M.,

Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreene,D., Fartan,D.,

Ferrera,S., Frisze,E., Galle,R.F., Gay,N.S., George,R.A.,

Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,

Idegam,C., Jallal,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,

McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,

Pacled,J., Pargay,V., Park,S., Patel,S., Pfeiffer,B.,

Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,

Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M.,

Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

Unpublished

2 (bases 1 to 170914)

Celniker,S.E., Agapayani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,

Butenhoff,C., Champs,M., Chavez,C., Chew,M., Ciesiolka,L.,

Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,

Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,

Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Madza,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,

Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,

Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.

Direct Submission

Submitted (24-SEP-1999) Drosophila Genome Center, Lawrence Berkeley

Laboratory, MS 54-121, Berkeley, CA 94720, USA

On May 12, 2001 this sequence version replaced gi:13374645.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome

shotgun and from subclones of this BAC and its neighboring clones.

For further information about this sequence, including its location

and relationship to other sequences, please visit our sequence

archive web site (<http://www.fruitfly.org/sequence/>) or send email

to hdg@fruitfly.berkeley.edu.

Location/Qualifiers

1. 170914

/organism="Drosophila melanogaster"

FEATURES

source

Location/Qualifiers

1. 87286 /organism="Drosophila melanogaster"

BASE COUNT

25135 a 18013 c 18299 g 25839 t

ORIGIN

Query Match

Best Local Similarity 56.2%; Score 26.4; DB 3; Length 170914;

Matches 27; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY

9 gcygacatawtygttcaytaagccgaaagcgtt 46

db 105904 GCTTACATATGTTCTCATTAAGCCTTAAGAACCT 105941

RESULT 14

LOCUS

AC010995 185810 bp DNA

INX 20-MAR-2001

DEFINITION Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone

AC010995 complete sequence.

ACCESSION

AC010995

VERSION

AC010995.11 GI:13384330

KEYWORDS

HTG.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 185810)

Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,

Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,

Rogers,Y., An,H., Baldwin,D., Banazon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M.,

Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreene,D., Fartan,D.,

Ferrera,S., Frisze,E., Galle,R.F., Gay,N.S., George,R.A.,

Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,

Idegam,C., Jallal,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,

McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,

Pacled,J., Pargay,V., Park,S., Patel,S., Pfeiffer,B.,

Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,

Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M.,

Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

Unpublished

2 (bases 1 to 185810)

Celniker,S.E., Agapayani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,

Butenhoff,C., Champs,M., Chavez,C., Chew,M., Ciesiolka,L.,

Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,

Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,

Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Madza,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,

Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,

Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.

Direct Submission

Submitted (29-SEP-1999) Drosophila Genome Center, Lawrence Berkeley

Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Mar 20, 2001 this sequence version replaced gi:6978354.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome

shotgun and from subclones of this BAC and its neighboring clones.

For further information about this sequence, including its location

and relationship to other sequences, please visit our sequence

archive web site (<http://www.fruitfly.org/sequence/>) or send email

to hdg@fruitfly.berkeley.edu.

Location/Qualifiers

1. 170914

/organism="Drosophila melanogaster"

FEATURES

source

Location/Qualifiers

1. 87286 /organism="Drosophila melanogaster"

BASE COUNT

25135 a 18013 c 18299 g 25839 t

ORIGIN

Query Match

Best Local Similarity 56.2%; Score 26.4; DB 3; Length 170914;

Matches 27; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY

9 gcygacatawtygttcaytaagccgaaagcgtt 46

db 105904 GCTTACATATGTTCTCATTAAGCCTTAAGAACCT 105941

RESULT 14

LOCUS

AC010995 185810 bp DNA

INX 20-MAR-2001

DEFINITION Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone

AC010995 complete sequence.

ACCESSION

AC010995

VERSION

AC010995.11 GI:13384330

KEYWORDS

HTG.

SOURCE

fruit fly.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 185810)

Celniker,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H.,

Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,

Rogers,Y., An,H., Baldwin,D., Banazon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M.,

Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dreene,D., Fartan,D.,

Ferrera,S., Frisze,E., Galle,R.F., Gay,N.S., George,R.A.,

Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,

Idegam,C., Jallal,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,

McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,

Pacled,J., Pargay,V., Park,S., Patel,S., Pfeiffer,B.,

Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,

Stapleton,M., Strong,R., Svirkas,R., Tector,C., Williams,S.M.,

Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.

Unpublished

2 (bases 1 to 185810)

Celniker,S.E., Agapayani,A., Arcaina,T.T., Baxter,E., Blazek,R.G.,

Butenhoff,C., Champs,M., Chavez,C., Chew,M., Ciesiolka,L.,

Doyle,C.M., Fartan,D.E., Galle,R., George,R.A., Harris,N.L.,

Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,

Kim,E., Lee,B., Lewis,S., Li,P., Lomolan,M.A., Madza,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacled,J.M., Park,S.,

Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,

Svirkas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.

Direct Submission

Submitted (29-SEP-1999) Drosophila Genome Center, Lawrence Berkeley

Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Mar 20, 2001 this sequence version replaced gi:6978354.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome

shotgun and from subclones of this BAC and its neighboring clones.

For further information about this sequence, including its location

and relationship to other sequences, please visit our sequence

archive web site (<http://www.fruitfly.org/sequence/>) or send email

to hdg@fruitfly.berkeley.edu.

Location/Qualifiers

1. 170914

/organism="Drosophila melanogaster"

FEATURES

source

Location/Qualifiers

1. 87286 /organism="Drosophila melanogaster"

BASE COUNT

25135 a 18013 c 18299 g 25839 t

ORIGIN

Query Match

Best Local Similarity 56.2%; Score 26.4; DB 3; Length 170914;

Matches 27; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY

9 gcygacatawtygttcaytaagccgaaagcgtt 46

db 105904 GCTTACATATGTTCTCATTAAGCCTTAAGAACCT 105941

RESULT 14

LOCUS

AC010995 185810 bp DNA

INX 20-MAR-2001

DEFINITION Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone

AC010995 complete sequence.

ACCESSION

AC010995

VERSION

AC010995.11 GI:13384330

KEYWORDS

HTG.

SOURCE

to bdpdefruitfly/berkeley.edu.

Location/Qualifiers

1.185810

/organism="Drosophila melanogaster"

/strain="y; cn bw sp"

/db_xref="taxon:7227"

/chromosome="X"

/map="12A-12A"

/clone="BACR30C16 (DB80)"

/clone_1ib="Rpci-98 (Roswell Park Cancer Institute

Drosophila melanogaster BAC library, partial EcORI in

PBACE3.6)"

BASE COUNT 54601 a 38842 c 38347 g 54020 t

ORIGIN

Query Match

Best Local Similarity

Matches

27; Conservative

5; Mismatches

6; Indels

0; Gaps

0;

RESULT 15

AE003492

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., and Venter J.C. The genome sequence of *Drosophila melanogaster* Science 287 (5461), 2185-2195 (2000)

2 (bases 1 to 298839)

Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., and Venter J.C. Direct Submission Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

On Oct 9, 2000 this sequence version replaced gi:7292850.

location/Qualifiers

1.298839

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/chromosome="X"

complement(427..>1038)

/gene="CG15742"

/product="CT35990"

/db_xref="FLYBASE:FBan0015742"

/db_xref="FLYBASE:FBgn0030462"

/evidence=not_experimental

complement(427..>1038)

/gene="CG15742"

/note="CG15742 gene product"

/codon_start=1

/db_xref="FLYBASE:FBan0015742"

/db_xref="FLYBASE:FBgn0030462"

/evidence=not_experimental

/protein_id="AAFA8243.1"

/db_xref="GI:7292851"

/translation="MSKRCNKKIDHCLGAVGKIPPRYSRVNADGVITPDKYSR/PNFATRIYVPAELPIDLAELIMDKQILRNKCIASSWVPMKICOGYGADETFIMDSAMKLOALIEELTNLDQKAVEVSGRLAVQQLDKRQIOMUNLNG/RNOLDFRARDISLPYPAIDARMSISKYVINMRTWRG"

join(1845..2057,2139..2231,2334..2524,2598..4226,4308..4466,4535..5296)

/gene="comt"

/note="Nucleotide sequence of the Celera sequence differs from the published sequence for this transcript."

/product="CT41661"

/db_xref="FLYBASE:FBan001618"

/db_xref="FLYBASE:FBgn0000346"

<1845..>5296

/gene="comt"

/note="CG1618"

/map="11E1-11E1"

/db_xref="FLYBASE:FBan001618"

/db_xref="FLYBASE:FBgn0000346"

join(2049..2057,2139..2231,2334..2524,2598..4226,4308..4466,4535..4667)

/gene="comt"

/note="comt gene product; Nucleotide sequence of the Celera sequence differs from the published sequence for this transcript"

/codon_start=1

/db_xref="FLYBASE:FBan001618"

/db_xref="FLYBASE:FBgn0000346"

/protein_id="AAFA8244.1"

/db_xref="GI:7292852"

/translation="MAYTIKATKPTDELSTINRAIVNVDPEIKYQHFITALEKT/VVPGYGFESIVGRKAMVINSINQELVPRPSSDVITCVSEFDFLOKTVSOE/PYDSOMKKEFMOPAGMALVGSILVFNFKKILGLAVSLKIDPKSLGSGKDA/MANVFGILNINVAQVPEKAKNSISLNQSGKGVQSIINPMDFGKKGIGLDKE/FSITRRFAASVPPPELVQDLGKHKXGILLYLPPTGKTIAMARQIGTMINAREPKI

CDS

gene

mRNA

mRNA

gene

CDS

Sat Feb 9 11:52:28 2002

us-09-548-998a-8.rge

Run on: February 9, 2002, 10:47:48 ; Search time 100.64 Seconds

(without alignments)
400.380 Million cell updates/sec

Title: US-09-548-998A-8
Perfect score: 47

Sequence: 1 tgtgtrtggcgcacatawgt.....ytaaagccgaaqqvqcltq 47

Scoring table: IDENTITY_NUC

Searched: 930621 sets, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Maximum DB seq length:	2000000000
------------------------	------------

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1:	/SID22_gcgdata/geneseq/geneseq/NAL1980.DAT *
2:	/SID22_gcgdata/geneseq/geneseq/NAL1981.DAT *
3:	/SID22_gcgdata/geneseq/geneseq/NAL1982.DAT *
4:	/SID22_gcgdata/geneseq/geneseq/NAL1983.DAT *
5:	/SID22_gcgdata/geneseq/geneseq/NAL1984.DAT *
6:	/SID22_gcgdata/geneseq/geneseq/NAL1985.DAT *
7:	/SID22_gcgdata/geneseq/geneseq/NAL1986.DAT *
8:	/SID22_gcgdata/geneseq/geneseq/NAL1987.DAT *
9:	/SID22_gcgdata/geneseq/geneseq/NAL1988.DAT *
10:	/SID22_gcgdata/geneseq/geneseq/NAL1989.DAT *
11:	/SID22_gcgdata/geneseq/geneseq/NAL1990.DAT *
12:	/SID22_gcgdata/geneseq/geneseq/NAL1991.DAT *
13:	/SID22_gcgdata/geneseq/geneseq/NAL1992.DAT *
14:	/SID22_gcgdata/geneseq/geneseq/NAL1993.DAT *
15:	/SID22_gcgdata/geneseq/geneseq/NAL1994.DAT *
16:	/SID22_gcgdata/geneseq/geneseq/NAL1995.DAT *
17:	/SID22_gcgdata/geneseq/geneseq/NAL1996.DAT *
18:	/SID22_gcgdata/geneseq/geneseq/NAL1997.DAT *
19:	/SID22_gcgdata/geneseq/geneseq/NAL1998.DAT *
20:	/SID22_gcgdata/geneseq/geneseq/NAL1999.DAT *
21:	/SID22_gcgdata/geneseq/geneseq/NAL2000.DAT *
22:	/SID22_gcgdata/geneseq/geneseq/NAL2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.2	94.0	47	22	AAC62245	A 16S rDNA sequen
2	44.2	94.0	121.2	22	AAC62239	A 16S rDNA sequen
3	44.2	94.0	121.2	22	AAC62241	A 16S rDNA sequen
4	44.2	94.0	121.2	22	AAC62242	A 16S rDNA sequen
5	44.2	94.0	121.2	22	AAC62243	A 16S rDNA sequen
6	44.2	94.0	133.5	22	AAC62240	A 16S rDNA sequen
7	44.2	94.0	144.3	22	AAC62244	A 16S rDNA sequen
8	21	44.7	146.8	22	AAH74593	Nucleotide sequen
c 9	21	44.7	811.5	22	AAFS4868	Genomic nucleotide
c 10	20.8	44.3	56.5	22	AAH06050	Human cDNA clone
11	20.8	44.3	156.0	22	AAF88043	Human Trp2 encodin

12	20.8	44.3	2291	18	AAAT91957	The novel tyrosin
C 13	20.6	43.8	1491	22	AAAT45594	Synthetic nucleot
C 14	20.6	43.8	3300	22	AAAT58569	Human polynucleot
C 15	20.6	43.8	3316	22	AAAT50355	Human polynucleot
C 16	20.6	43.8	3363	20	AAAT67555	Candida RML codin
C 17	20.4	43.4	147	21	AAAT659054	Bacteriophage 44A
C 18	20.4	43.4	478	21	AAAT16377	Probe #6510 for g
C 19	20.4	43.4	478	22	AAAT39032	Probe #7718 used
C 20	20.4	43.4	787	21	AAAT43633	Human secreted pr
C 21	20.4	43.4	1464	21	AAAT69017	Bacteriophage 44A
C 22	20.4	43.4	1666	21	AAAT69013	Bacteriophage 44A
C 23	20.2	43.0	1293	21	AAAT88420	Human Est-derived
C 24	20.2	43.0	1937	21	AAAT7445	Macaque TANGO 232
C 25	20.2	43.0	4594	18	AAAT4511	Staphylococcus aut
C 26	20.2	43.0	10207	15	AAAT055134	Staphylococcus aut
C 27	20	42.6	437	20	AAAT90537	Est. locus BK498
C 28	20	42.6	3444	22	AAAT83798	Arididopsis thal
C 29	19.8	42.1	47	21	AAAT44357	Human preproinsulin
C 30	19.8	42.1	480	22	AAAT04668	Probe #4859 used t
C 31	19.8	42.1	8355	17	AAAT35524	Human Interferon- β
C 32	19.6	41.7	375	21	AAAT60386	A diacylglycerol a
C 33	19.6	41.7	636	22	AAAT03401	Human cDNA clone (
C 34	19.6	41.7	689	21	AAAT02024	Human colon cancer
C 35	19.6	41.7	689	21	AAAT02024	Human colonic cancer
C 36	19.6	41.7	1476	22	AAAT15552	Human stromalin-1
C 37	19.6	41.7	4150	16	AAAT08232	Human cDNA sequenc
C 38	19.6	41.7	319608	22	AAAT51601	Murine stromalin-1
C 39	19.6	41.7	319608	22	AAAT09301	Human chromosome 1
C 40	19.6	41.3	580073	18	AAAT58840	Human schizoprenia
C 41	19.4	41.3	221	20	AAAT31851	Mycoplasma genital
C 42	19.4	41.3	466	21	AAAT15984	S. aureus polypept
C 43	19.4	41.3	557	21	AAAT57552	Human prostate can
C 44	19.4	41.3	558	21	AAAT56216	Pinus radiata tran
C 45	19.4	41.3	714	22	AAAT67447	Pinus radiata tran
	19.4	41.3	762	19	AAAT59777	C. glutamicum codin
						Human secreted pro

ALIGNMENTS

RESULT	1
AC662245	
ID	MAC62245 standard; DNA; 47 BP.
AC	MAC62245;
XX	
DT	19-MAR-2001 (first entry)
XX	
DE	A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX	
KW	16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride
KW	carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
XX	trichloroethane; dichloroethylene; chlorinating bacteria; ss.
OS	Synthetic.
XX	
OS	Dehalococcioides ethenogenes.
XX	
PN	MO200063443-A2.
XX	
PD	26-OCT-2000.
XX	
PF	13-APR-2000; 2000WO-US09883.
XX	
PR	15-APR-1999; 99US-0129511.
XX	
PA	(DUPO) DU POINT DE NEMOURS & CO E I.
XX	
PI	Hendrickson ER, Ebersole RC;
XX	
DR	WPI; 2001-024581/03.
XX	
PT	New 16S rDNA profile derived from Dehalococcioides ethenogenes and
PT	indicative of a dechlorinating bacterial strain -

PS Claim 1; Page 50; 55pp; English.

CC The present sequence represents a consensus 16S rDNA sequence, derived
CC from various strains of Dehalococcoides ethenogenes. The 16S rDNA
CC profile is linked to dechlorinating activity. Bacterial strain
CC comprising the 16S rDNA sequence of the invention are useful for the
CC dechlorination of chlorinated compounds such as carbon tetrachloride,
CC tetrachloroethane, chloroform, dichloroethane, trichloroethane,
CC dichloroethylene, vinyl chloride, and chloroformatics. The 16S rDNA
CC sequence is also useful for identification of new chlorinating bacteria,
CC as well as sub-typing strains of Dehalococcoides ethenogenes.

SO Sequence 47 BP; 9 A; 6 C; 15 G; 10 T; 7 other;

Query Match 94.0%; Score 44.2; DB 22; Length 47;
Best Local Similarity 100.0%; Pred. No. 1e-09; 0; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0;

1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 47
1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 47

RESULT 2

AAC62239
ID AAC62239 standard; DNA; 1212 BP.

AC AAC62239;

DT 19-MAR-2001 (first entry)

DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.

XX 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbon tetrachloride; tetrachloroethane; chloroform; dichloroethane;
KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.

OS Dehalococcoides ethenogenes.

PN WO200063443-A2.

PD 26-OCT-2000.

PF 13-APR-2000; 2000WO-US09883.

PR 15-APR-1999; 99US-0129511.

PS (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Hendrickson ER, Ebersole RC;

DR WPI; 2001-024581/03.

PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
PT indicative of a dechlorinating bacterial strain.

PS Claim 1; Page 47; 55pp; English.

CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain PL, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbon tetrachloride, tetrachloroethane, chloroform, dichloroethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.

SO Sequence 1212 BP; 321 A; 252 C; 369 G; 269 T; 1 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 85.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 47
1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 186

RESULT 3

AAC62241
ID AAC62241 standard; DNA; 1212 BP.

AC AAC62241;

DT 19-MAR-2001 (first entry)

DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.

XX 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbon tetrachloride; tetrachloroethane; chloroform; dichloroethane;
KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.

OS Dehalococcoides ethenogenes.

PN WO200063443-A2.

PD 26-OCT-2000.

PF 13-APR-2000; 2000WO-US09883.

PR 15-APR-1999; 99US-0129511.

PS (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Hendrickson ER, Ebersole RC;

DR WPI; 2001-024581/03.

PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
PT indicative of a dechlorinating bacterial strain.

PS Claim 1; Page 48; 55pp; English.

CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain DAB, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbon tetrachloride, tetrachloroethane, chloroform, dichloroethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.

SO Sequence 1212 BP; 318 A; 253 C; 372 G; 269 T; 0 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 85.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 47
1 tgggtggcgagacatawgttgctcaatgaagccgaaagggcttg 186

RESULT 4

AAC62242
ID AAC62242 standard; DNA; 1212 BP.

AC AAC62242;

XX

DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KW 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbonotetrachloride; tetrachloroethane; chloroform; dichloromethane;
XX trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.
XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Hendrickson ER, Ebersole RC;
XX WPI: 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
XX indicative of a dechlorinating bacterial strain -
XX
PS Claim 1; Page 48-49; 55pp; English.
XX
CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain PIN, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbonotetrachloride, tetrachloroethane, chloroform, dichloromethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.
XX
SQ Sequence 1212 BP; 320 A; 253 C; 371 G; 268 T; 0 other.

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 35.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgttgggagacatawgttggttcaytaagccggaagggctgtg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 tctgttgggagacataagttggttcactaaagccgtaagtgctgtg 186

RESULT 5
AAC62243
ID AAC62243 standard; DNA; 1212 BP.
XX
AC AAC62243;
XX
DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KW 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbonotetrachloride; tetrachloroethane; chloroform; dichloromethane;
XX trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.
XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX

XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Hendrickson ER, Ebersole RC;
XX WPI: 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
XX indicative of a dechlorinating bacterial strain -
XX
PS Claim 1; Page 49; 55pp; English.
XX
CC The present sequence represents the 16S rDNA profile of Dehalococcoides
CC ethenogenes strain DL, isolated from soil surrounding an industrial
CC site. The 16S rDNA profile is linked to dechlorinating activity.
CC Bacterial strain comprising the 16S rDNA sequence of the invention are
CC useful for the dechlorination of chlorinated compounds such as
CC carbonotetrachloride, tetrachloroethane, chloroform, dichloromethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
CC ethenogenes.
XX
SQ Sequence 1212 BP; 320 A; 255 C; 371 G; 266 T; 0 other.

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 85.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgttgggagacatawgttggttcaytaagccggaagggctgtg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 tctgttgggagacataagttggttcactaaagccgtaagtgctgtg 186

RESULT 6
AAC62240
ID AAC62240 standard; DNA; 1335 BP.
XX
AC AAC62240;
XX
DT 19-MAR-2001 (first entry)
XX
DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.
XX
KW 16S rDNA: dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbonotetrachloride; tetrachloroethane; chloroform; dichloromethane;
XX trichloroethane; dichloroethylene; chlorinating bacteria; ss.
XX
OS Dehalococcoides ethenogenes.
XX
PN WO200063443-A2.
XX
PD 26-OCT-2000.
XX
PF 13-APR-2000; 2000WO-US09883.
XX
PR 15-APR-1999; 99US-0129511.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Hendrickson ER, Ebersole RC;
XX WPI: 2001-024581/03.
XX
PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
XX indicative of a dechlorinating bacterial strain -
XX
PS Claim 1; Page 47-48; 55pp; English.
XX
CC The present sequence represents the 16S rDNA profile of Dehalococcoides

CC ethenogenes strain STF, isolated from soil surrounding an industrial
 CC site. The 16S rDNA profile is linked to dechlorinating activity.
 CC Bacterial strain comprising the 16S rDNA sequence of the invention are
 CC useful for the dechlorination of chlorinated compounds such as
 CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
 CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
 CC The 16S rDNA sequence is also useful for identification of new
 CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
 CC ethenogenes.
 CC
 SQ Sequence 1335 BP; 348 A; 288 C; 403 G; 296 T; 0 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1335;
 Best Local Similarity 85.1%; Pred. No. 2.8e-09;
 Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgtgagcgcacatawgttgctcaytaagccgaaagggcttg 47
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 140 tgtgtgtgagcgcacataagtgctcactaagccgaaagggcttg 186

RESULT 7
 AAC62244
 ID AAC62244 standard; DNA; 1443 BP.

XX AAC62244;

DT 19-MAR-2001 (first entry)

DE A 16S rDNA sequence indicative of a chlorinating bacterial strain.

XX 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
 KW carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
 KM trichloroethane; dichloroethylene; chlorinating bacteria; ss.

XX Dehalococcoides ethenogenes.

XX WO200063443-A2.

XX 26-OCT-2000.

XX 13-APR-2000; 2000WO-US09883.

XX 15-APR-1999; 99US-0129511.

PA (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Hendrickson ER, Ebersole RC;

XX WPI; 2001-024581/03.

PT New 16S rDNA profile derived from Dehalococcoides ethenogenes and
 PT indicative of a dechlorinating bacterial strain -

XX Example 2; Page 49-50; 55pp; English.

XX The present sequence represents the 16S rDNA profile of Dehalococcoides
 CC ethenogenes strain 195. The 16S rDNA profile is linked to dechlorinating
 CC activity. Bacterial strain comprising the 16S rDNA sequence of the
 CC invention are useful for the dechlorination of chlorinated compounds such
 CC as carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
 CC trichloroethane, dichloroethylene, vinyl chloride, and chloroformatics.
 CC The 16S rDNA sequence is also useful for identification of new
 CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
 CC ethenogenes.
 CC
 SQ Sequence 1443 BP; 379 A; 306 C; 443 G; 314 T; 1 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1443;
 Best Local Similarity 85.1%; Pred. No. 2.8e-09;
 Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgtgagcgcacatawgttgctcaytaagccgaaagggcttg 47
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 140 tgtgtgtgagcgcacataagtgctcactaagccgaaagggcttg 186

RESULT 8
 AAH74593/C
 ID AAH74593 standard; DNA; 1488 BP.

XX AAH74593;

DT 15-OCT-2001 (first entry)

DE Nucleotide sequence of a hydantoinase enzyme.

XX Methionine gamma-lyase; mda gene; free folding energy; gene shuffling;
 KW directed evolution; molecular breeding; hydantoinase; ss.

XX Pseudomonas putida.

XX Key location/Qualifiers
 FH 1.1488
 FT /*tag= a
 FT CDS /product= "hydantoinase"

XX WO20015342-A2.

XX 02-AUG-2001.

XX 31-JAN-2001; 2001WO-US03186.

XX 31-JAN-2000; 2000US-0494921.

XX 08-DEC-2000; 2000US-0734237.

XX (BIOC-) BIOCATALYTICS INC.

XX Rozzell DJ, Bui P, Hua L;

XX WPI; 2001-483235/52.

XX P-PADB; AAC63575.

PT Designing synthetic nucleic acid sequences for improved amplification,
 PT expression in host cell, by comparing free energy of folding of a
 PT starting polynucleotide and a modified polynucleotide having a codon
 PT replacement -

XX Example 9; Page 91-92; 117pp; English.

XX The present sequence encodes a hydantoinase enzyme. The polynucleotide
 CC was modified using the method of the invention. The
 CC specification describes a method for designing a synthetic
 CC polynucleotide. The method comprises providing a starting polynucleotide,
 CC determining the predicted free energy of folding per base of the
 CC polynucleotide, modifying the polynucleotide by replacing a codon with
 CC a different codon to provide a modified polynucleotide, determining free
 CC energy of folding per base of the modified polynucleotide, and
 CC comparing this with that of the original polynucleotide. The method is
 CC useful for developing nucleic acid sequences that enhance expression of
 CC the encoded protein in a heterologous host. The design and preparation
 CC of the synthetic genes are used in application of gene shuffling,
 CC directed evolution and molecular breeding methods. The method allows
 CC expression of genes from various organisms such as mammals, plants,
 CC yeast, fungi and bacteria in prokaryotic hosts, such as Escherichia coli
 CC and eukaryotic hosts at commercially viable levels, in particular
 CC proteins with low yield such as methionine gamma-lyase from P. putida.

XX Sequence 1488 BP; 285 A; 502 C; 456 G; 245 T; 0 other;

Query Match 44.7%; Score 21; DB 22; Length 1488;
 Best Local Similarity 57.4%; Pred. No. 26;
 Matches 27; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

KW anyotropic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.
2000US-0552317

09-JUL-2000; 2000US-0598042.
10-JUL-2000; 2000US-0620312.

PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036
29-NOV-2000; 2000US-0727344

XX
DA (HYSE-) HYSEO INC.

XX	Tand VT.	Iju C,	Asundi V,
PT			

PI	Wang J,	Wang Z,	Wentman T
PI	Zhao OA,	Zhou P,	Goodrich

XX
DR WPT: 2001-442253/47.

DR P-PSDB; AAM41199.
XX

PT Novel nucleic acid
PT such as central n

XX
PS Claim 1; SEQ ID NO 4344; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and

the encoded polypeptides (AAW38042-AAW42137) which nucleotides are useful immunosuppressant and cytostatic activity. The polynucleotides or polynucleotid In gene therapy. A composition containing a polypeptide or polypeptides of the system, such as peripheral nervous injuries, peripheral neuropathy and Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 3316 BP; 1013 A; 595 C; 720 G; 983 T; 5 other, SQ

43 89. score 20.6: DB 22: Length 3316

Matches	23;	Conservative	5;	Mismatches	9;	Indels	0;	Gaps	0
23	23	5	5	9	9	0	0	0	0

QY 10 cygacatawgtggttcaytaagccggaagygctt 46
 ||||| : |||||:||||| :||| : |||||
 Db 1523 ctgcatcttttaagttcattaagcagcaaatctct 1555

```
Search completed: February 9, 2002, 11:33:39
Job time: 2751 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 38.58 Seconds

(without alignments)
275,906 Million cell updates/sec

Title: US-09-548-998a-8

Perfect score: 1 tggtygggcyacatawgt.....ytaaagccgaagygcttg 47

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCtUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	20.8	44.3	2291 2	US-08-725-736D-1
2	20.8	44.3	2291 3	US-09-162-368B-1
3	20.8	44.3	2291 3	US-09-161-877B-1
4	20.6	43.8	3463 4	US-09-189-462-3
5	20.2	43.0	10207 1	US-08-920-812-2
6	20.2	43.0	10207 1	US-08-920-827-2
7	20.2	43.0	10207 1	US-08-921-177-2
8	20.2	43.0	10207 1	US-08-362-577C-2
9	20.2	43.0	10207 2	US-08-920-828-2
10	20.2	42.6	269 4	US-09-060-756-144
11	19.8	42.1	8355 4	US-08-406-030A-23
12	19.4	41.3	2499 4	US-09-205-283-1
13	19.4	41.3	2499 4	US-09-205-283-1
14	19.4	41.3	2520 4	US-08-968-563-10
15	19.4	41.3	2520 3	US-08-969-683A-10
16	19.4	41.3	19227 3	US-09-090-793-13
17	19.4	41.3	40138 3	US-09-090-793-12
18	19.2	40.9	589 4	US-09-385-982-454
19	19.2	40.9	1764 4	US-08-980-832-36
20	19.2	40.9	2033 2	US-08-417-210A-136
21	19.2	40.9	2033 2	US-08-417-210A-139
22	19.2	40.9	2063 2	US-08-417-210A-142
23	19.2	40.9	3040 1	US-08-446-794A-1
24	19.2	40.9	3040 1	US-08-750-007-2
25	19.2	40.9	3040 1	US-08-945-024-1
26	19.2	40.9	3808 2	US-08-417-210A-79
27	19.2	40.9	15328 2	US-08-888-497-33

28	19.2	40.9	15328 5	PCT-US94-07926-33	Sequence 33, Appl
29	19	40.4	1688 2	US-08-793-410-1	Sequence 1, Appl
30	19	40.4	1688 2	US-08-793-410-27	Sequence 27, Appl
31	19	40.4	3515 1	US-08-596-985-1	Sequence 1, Appl
32	19	40.4	7333 4	US-08-766-528-2	Sequence 1, Appl
33	19	40.4	8060 4	US-08-766-528-3	Sequence 2, Appl
34	19	40.4	8132 4	US-08-353-700-2	Sequence 3, Appl
35	19	40.4	10136 1	US-08-353-700-2	Sequence 2, Appl
36	19	40.4	10136 5	PCT-US95-16216-2	Sequence 2, Appl
37	18.8	40.0	387 1	US-08-423-383-2	Sequence 2, Appl
38	18.8	40.0	387 2	US-08-437-353A-2	Sequence 2, Appl
39	18.8	40.0	1164 6	5240849-4	Sequence 2, Appl
40	18.8	40.0	2749 6	5240849-1	Patent No. 5240849
41	18.6	39.6	1318 4	US-08-986-304-1	Sequence 1, Appl
42	18.6	39.6	3240 1	US-08-294-189-2	Sequence 1, Appl
43	18.6	39.6	5494 4	US-08-910-864-10	Sequence 2, Appl
44	18.6	39.6	5510 1	US-08-123-161A-7	Sequence 7, Appl
45	18.6	39.6	5510 1	US-08-483-278-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-725-736D-1
Sequence 1, Application US/08725736D
Patent No. 5831016
GENERAL INFORMATION:
APPLICANT: MANG, R.F.; ROSENBERG, S. A.
TITLE OF INVENTION: IDENTIFICATION OF TRP-2 AS
TITLE OF INVENTION: A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & PINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,736D
FILING DATE: 04-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,602
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36,434
REFERENCE/DOCKET NUMBER: 2026-4243
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2291
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: CDNA
DESCRIPTION: CDNA
FEATURE:
NAME/KEY: TRP-2 GENE
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:

DESCRIPTION: CDNA
FEATURE:
NAME/KEY: TRP-2 GENE
LOCATION:
IDENTIFICATION METHOD:
PUBLICATION INFORMATION:
AUTHORS: YOKAYAMA, ET AL.
TITLE: MOLECULAR CLONING AND FUNCTIONAL
ANALYSIS OF A CDNA CODING FOR HUMAN
DOPACHROME TATTOERASE/TYROSINASE
TITLE: RELATED PROTEIN-2.
JOURNAL: BIOCHIM. BIOPHYS. ACTA.
VOLUME: 1217
ISSUE:
PAGES: 317-321
DATE: 1994
US-09-161-877B-1

Query Match 44.3%; Score 20.8; DB 3; Length 2291;
Best Local Similarity 59.5%; Pred. No. 7.8;
Matches 25; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 2 ggtgtggcgcgcacatagtggttcacaaagccgaaaggg 43
DB 1507 GGTGATGACCTTCATATTGTCATCTTCCTGAAACGG 1548

RESULT 4
US-09-189-462-3/C
Sequence 3, Application US/09189462
Patent No. 630302
GENERAL INFORMATION:
APPLICANT: Rupp, Steffan
APPLICANT: Robertson, Laura
APPLICANT: Summers, Eric F.
APPLICANT: Hecht, Peter
APPLICANT: Roberts, Radclyffe
APPLICANT: Madhani, Hiren
APPLICANT: Styles, Cora Ann
APPLICANT: Lo, Hsiu-Jung
APPLICANT: Sherman, Amir
APPLICANT: Call, Brian
TITLE OF INVENTION: REGULATION OF FUNGAL GENE EXPRESSION
FILE REFERENCE: 50078/003002
CURRENT APPLICATION NUMBER: US/09/189,462
CURRENT FILING DATE: 1998-11-10
EARLIER APPLICATION NUMBER: 60/066,129
EARLIER FILING DATE: 1997-11-19
EARLIER APPLICATION NUMBER: 60/066,308
EARLIER FILING DATE: 1997-11-21
EARLIER APPLICATION NUMBER: 60/066,462
EARLIER FILING DATE: 1997-11-24
EARLIER APPLICATION NUMBER: 60/078,610
EARLIER FILING DATE: 1998-03-19
EARLIER APPLICATION NUMBER: 60/094,523
EARLIER FILING DATE: 1998-07-29
NUMBER OF SEQ ID NOS: 67
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 3463
TYPE: DNA
ORGANISM: Candida albicans
US-09-189-462-3

Query Match 43.8%; Score 20.6; DB 4; Length 3463;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 26; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 tgggtggcgcgcacatagtggttcacaaagccgaaagggcttg 47

DB 787 TGTGTTGTGATGACATAGTTGTTGTTGTAATAATAGATGTTG 741

RESULT 5
US-08-920-812-2/C
Sequence 2, Application US/08920812
Patent No. 5763188
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,812
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-920-812-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 10 cygacatawtggttcacaaagccgaaagggctt 46
DB 6253 CTGACATACATCTTCACATATACCGCAGAGTACTT 6217

RESULT 6
US-08-920-827-2/C
Sequence 2, Application US/08920827
Patent No. 5770375
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji

TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-920-827-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

10 cygacatawgtggtcaytaagccgyaagggctt 46
1:|||||: : |||: | |||: |||
Db 6253 CTGACATACATCTTCCACATACCGCAGGTACTT 6217

RESULT 7
US-08-921-177-2/c
Sequence 2, Application US/08921177
Patent No. 5798211
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,177
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-921-177-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

10 cygacatawgtggtcaytaagccgyaagggctt 46
1:|||||: : |||: | |||: |||
Db 6253 CTGACATACATCTTCCACATACCGCAGGTACTT 6217

RESULT 8
US-08-362-577C-2/c
Sequence 2, Application US/08362577C
Patent No. 5807673
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,577C
FILING DATE: 27-MAR-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-362-577C-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 10 cygacatawtyggttcaytaagccgyaagygctt 46
1:|||||: : |||: | ||: |||: |||
Db 6253 CTGACATAACATCTTCACATACGCCGCAAGTACTT 6217

RESULT 9

US-08-920-828-2/C
Sequence 2, Application US/08920828
Patent No. 5853998
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,828
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6500
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-920-828-2

Query Match 43.0%; Score 20.2; DB 2; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 10 cygacatawtyggttcaytaagccgyaagygctt 46
1:|||||: : |||: | ||: |||: |||
Db 6253 CTGACATAACATCTTCACATACGCCGCAAGTACTT 6217

RESULT 10

US-09-060-756-144/C
Sequence 144, Application US/09060756
Patent No. 6183957
GENERAL INFORMATION:
APPLICANT: Cole, Stewart
APPLICANT: Buchrieser-Brosch, Roland
APPLICANT: Gordon, Stephen
APPLICANT: Billault, Alain
TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
FILE REFERENCE: 3495-0169
CURRENT APPLICATION NUMBER: US/09/060,756
CURRENT FILING DATE: 1998-04-16
NUMBER OF SEQ ID NOS: 743
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 144
LENGTH: 269
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
NAME/KEY: unsure
LOCATION: (various positions within the sequence)
OTHER INFORMATION: applicants are uncertain of bases designated as "n"
US-09-060-756-144

Query Match 42.6%; Score 20; DB 4; Length 269;
Best Local Similarity 59.1%; Pred. No. 8.7;
Matches 26; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 4 grtggccygcatawtyggttcaytaagccgyaagygctt 47
1:|||||: : |||: | ||: |||: |||
Db 243 GGTGCCCGGACGAGAAAGCTTCAATTGCGCAAGGTCTTG 200

RESULT 11

US-08-406-030A-23
Sequence 23, Application US/08406030A
Patent No. 6270989
GENERAL INFORMATION:
APPLICANT: Treco, Douglas A.
APPLICANT: Heartlein, Michael W.
APPLICANT: Haughe, Brian M.
APPLICANT: Selden, Richard F.
TITLE OF INVENTION: Protein Production and Delivery
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,030A
FILING DATE: 17-MAR-1995
CLASSIFICATION: 435

RESULT 12
 US-09-205-283-1/C
 : Sequence 1, Application US/09205283
 : Patent No. 6183973
 : GENERAL INFORMATION:
 : APPLICANT: Wright C, Anita
 : APPLICANT: Powell L, Jan
 : APPLICANT: Morris J, Glenn
 : APPLICANT: University Of Maryland Biotechnology Institute
 : TITLE OF INVENTION: V. Vulnificus Molecular Probes Antibodies, and Proteins
 : FILE REFERENCE: 000433-00010
 : CURRENT APPLICATION NUMBER: US/09/205, 283
 : CURRENT FILING DATE: 1998-12-04
 : EARLIER APPLICATION NUMBER: PCT/US98/12467
 : EARLIER FILING DATE: 1998-06-19
 : NUMBER OF SEQ ID NOS: 14
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO. 1
 : LENGTH: 2499
 : TYPE: DNA
 : ORGANISM: Vibrio vulnificus
 : US-09-205-283-1

Query Match	41.3%;	Score 19.4;	DB 4;	Length 2499;
Best Local Similarity	53.5%;	Pred. No. 32;		

Query Match	Similarity	41.3%	Score 19.4	DB 4	Length 2499
Best Local	Similarity	53.5%	Pred. No. 32		
Matches	23	Conservative	7	Mismatches	13
				Indels	0
				Gaps	0
QY	5	rtggagcygacatawtyggttcacaytaaaagcgyaaagygcttg	47		
Db	2274	gttaagacacacttcagttgagttctcatttaaaagccatnaatcgttgc	2232		

US-08L-968-563-10/c
Sequence 10, Application US/08968563
Patent No. 6013494

GENERAL INFORMATION:

APPLICANT: CHARLES E. NAKAMURA
APPLICANT: ATTHONY A. GATEMBY
APPLICANT: AMY (KUANG-HUA) HSU
APPLICANT: RICHARD D. LA REAU
APPLICANT: SHARON L. HAYNIE
APPLICANT: MARIA DIAZ-TORRES
APPLICANT: DONALD E. TRIMBUR
APPLICANT: GREGORY M. WHITED
APPLICANT: VASANTHA NAGARAJAN
APPLICANT: MARK S. PAVNE
APPLICANT: STEPHEN K. PICATAGGIO
APPLICANT: RAMESH V. NAIR
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898

ADDRESSEE: GENENCOR INTERNATIONAL, INC.
STREET: 4 CAMBRIDGE PLACE
CITY: ROCHESTER
STATE: NEW YORK
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
SOFTWARE: MICROSOFT WORD VERSION 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,563
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/030,601
FILING DATE: NOVEMBER 13, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9982
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT1
US-08-968-563-10

Query Match 41.3%; Score 19.4; DB 3; Length 2520;
Best Local Similarity 59.0%; Pred. No. 32;
Matches 23; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 tgtgtggygcgcacatagtggttcaytaagccgaa 39
|||:|||||:|||||:|||||:|||||:|
Db 95 TGTATGTTTACATACGTTGTTAATAAAGTCGTTA 57

RESULT 15
US-08-969-683A-10/c
Sequence 10, Application US/08969683A
Patent No. 6136576
GENERAL INFORMATION:
APPLICANT: GENENCOR INTERNATIONAL, INC.
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genencor International, Inc.
STREET: 4 Cambridge Place
CITY: Rochester
STATE: NY
COUNTRY: U.S.A
ZIP: 14618
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/969,683A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/20873
FILING DATE: 13-NOV-1997
APPLICATION NUMBER: 60/030,601
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gialster, Debra
REGISTRATION NUMBER: 33,888
REFERENCE/DOCKET NUMBER: GC 369-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-864-7620
TELEFAX: 650-845-6504
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2520 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: GUT1
US-08-969-683A-10

Query Match 41.3%; Score 19.4; DB 3; Length 2520;
Best Local Similarity 59.0%; Pred. No. 32;
Matches 23; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 tgtgtggygcgcacatagtggttcaytaagccgaa 39
|||:|||||:|||||:|||||:|||||:|
Db 95 TGTATGTTTACATACGTTGTTAATAAAGTCGTTA 57

Search completed: February 9, 2002, 10:48:51
Job time: 63 sec

Sat Feb 9 11:52:29 2002

us-09-548-998a-8.rni

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 1178.28 Seconds

(without alignments)
428.635 Million cell updates/sec

Title: US-09-548-998a-8

Perfect score: 1 tgyrtgcygcacatawtg.....ytaaagccygaagygcttg 47

Sequence:

Scoring table: IDENTITY-NMC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537288281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthmm:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estdv:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hmm:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	24.8	52.8	869	13	AZ136776 SP_0175_A
C 2	24.4	51.9	510	13	A0872190 nbe0047K
C 3	24	51.1	438	10	A1777700 EST258495
C 4	23.8	50.6	564	11	B1181099 UNL-P-PN-
C 5	23.8	50.6	564	11	B1181099 UNL-P-PN-
C 6	23.2	49.4	1028	13	CNS02BOM
C 7	23	48.9	469	13	A0761792 Tetradon
C 8	23	48.9	689	13	AQ312407 RPT11-10
C 9	23	48.9	835	11	BG585334 EST487098
C 10	22.6	48.5	267	13	TA35BE08P
C 11	22.6	48.1	520	10	AA355599 cp26f06.s
C 12	22.6	48.1	565	11	BG948475 IPL10_BI

C 13	22.6	48.1	598	13	A0699189
C 14	22.6	48.1	1008	11	BG331740
C 15	22.4	47.7	354	11	BG089520
C 16	22.4	47.7	422	10	AW124911
C 17	22.4	47.7	502	10	BE358808
C 18	22.4	47.7	525	13	TA288D10P
C 19	22.4	47.7	575	11	B1183356
C 20	22.4	47.7	683	10	BE650336
C 21	22.4	47.7	905	11	BE590336
C 22	22.2	47.2	403	10	AV630336
C 23	22.2	47.2	449	10	AW758129
C 24	22.2	47.2	449	10	AW640450
C 25	22.2	47.2	472	13	A2126430
C 26	22.2	47.2	489	13	A0868010
C 27	22.2	47.2	532	13	A0500765
C 28	22.2	47.2	659	10	AM963081
C 29	22.2	46.8	613	11	BG353081
C 30	22.2	46.8	642	13	A2730134
C 31	22.2	46.8	673	11	BG732137
C 32	22.2	46.8	1006	13	CNS06VY9
C 33	21.8	46.4	445	10	AA264746
C 34	21.8	46.4	457	10	AA951187
C 35	21.8	46.4	463	10	AA441387
C 36	21.8	46.4	532	10	AM531096
C 37	21.8	46.4	556	11	B1162978
C 38	21.8	46.4	578	13	A2102232
C 39	21.8	46.4	581	13	A2824456
C 40	21.8	46.4	586	13	TA388A06P
C 41	21.8	46.4	835	11	BG721676
C 42	21.8	46.4	870	11	BF537581
C 43	21.8	46.4	903	13	CNS06VY9
C 44	21.8	46.4	942	11	BF793898
C 45	21.8	46.4	952	11	BE86937

ALIGNMENTS

RESULT 1
AZ136776/c
LOCUS
DEFINITION
SP_0175.AL.F01.T7A Strongylocentrotus purpuratus, purple sea urchin clone genomic BAC library Strongylocentrotus purpuratus genomic clone Plate-175 Col-1 Row-K, DNA sequence.

ACCESSION
AZ136776
VERSION
AZ136776.1 GI:8288679
KEYWORDS
GSS.
SOURCE
Strongylocentrotus purpuratus.
ORGANISM
Strongylocentrotus purpuratus.
Eukaryote; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidae; Euechinoidae; Echinacea; Echinoida; Strongylocentrotidae; Strongylocentrotus.

REFERENCE
1 (bases 1 to 869)
Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R., Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and Hood,L.
A sea urchin genome project: Sequence scan, virtual map, and additional resources
Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)

TITLE
JOURNAL
MEDLINE
COMMENT
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 175 row: K column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 869.
Location/Qualifiers

source 1.869
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="Plate-175 Col-1 Row-K"
/clone_lib="Strongylocentrotus purpuratus, purple sea urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BAC3.6; BAC clones in E-Coli DH10B"

BASE COUNT 176 a 160 c 207 g 326 t
ORIGIN

Query Match 52.8%; Score 24.8; DB 13; Length 869;
Best Local Similarity 61.9%; Pred. No. 25;
Matches 26; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 1 tggtrggcgcacatawgttggttcaytaagccgaaaggy 42
|||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 239 TGAGTGGGCTGCATCAGAGGCTCCACCAAGTACGAGGC 198

RESULT 2
LOCUS A0872190 510 bp DNA GSS 03-NOV-1999
DEFINITION nbe0047K03f CUGI Rice BAC Library (EcoRI) Oryza sativa genomic clone nbe0047K03f, DNA sequence.
ACCESSION A0872190
VERSION A0872190.1 GI:6222641
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 510)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 42
High quality sequence stop: 202.
Location/Qualifiers

FEATURES
source 1.510
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbe0047K03f"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/tissue_type="leaf"
/lab_host="E. coli DH10B"
/note="Vector: pBACindigo; Site_1: EcoRI; Site_2: EcoRI; Rice is the most important food crop in the world. Half of the world population, especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 35,236 clones with an average insert size

of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

BASE COUNT 174 a 103 c 89 g 143 t 1 others
ORIGIN

Query Match 51.9%; Score 24.4; DB 13; Length 510;
Best Local Similarity 60.9%; Pred. No. 32;
Matches 28; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Oy 1 tggtrggcgcacatawgttggttcaytaagccgaaaggy 46
|||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 420 TGATTTGAGCTTTCTATGTTGATTTAAAGCTGAAGCGGCTT 375

RESULT 3
LOCUS A1777700 438 bp mRNA EST 18-MAY-2001
DEFINITION EST258495 tomato susceptible, Cornell Lycopersicon esculentum CDNA clone CLES2F21, mRNA sequence.
ACCESSION A1777700
VERSION A1777700.1 GI:5275657
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 438)
D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Renning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nieman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D., Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from Pseudomonas susceptible tomato Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
Location/Qualifiers

FEATURES
source 1.438
/organism="Lycopersicon esculentum"
/cultivar="Ril-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLES2F21"
/clone_lib="tomato susceptible, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; CLES - tomato Pseudomonas Susceptible EST library. directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT 131 a 87 c 109 g 111 t
ORIGIN

Query Match 51.1%; Score 24; DB 10; Length 438;
Best Local Similarity 61.4%; Pred. No. 45;
Matches 27; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Oy 1 tggtrggcgcacatawgttggttcaytaagccgaaaggy 44
|||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 345 TGGTGGAGCTTACATTCCTCCATTCACTAAGCAGCGAGTCC 388